

AD_____

Award Number: W81XWH-12-1-0550

TITLE: Early ICU Standardized Rehabilitation Therapy for the Critically Injured Burn Patient

PRINCIPAL INVESTIGATOR: Peter E. Morris, MD

CONTRACTING ORGANIZATION: Wake Forest University Health Sciences
Winston Salem, NC 27157

REPORT DATE: October 2014

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
<small>Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</small>					
1. REPORT DATE 17/10/2014		2. REPORT TYPE Annual		3. DATES COVERED 1 Oct 2013 – 30 Sep 2014	
4. TITLE AND SUBTITLE Early ICU Standardized Rehabilitation Therapy for the Critically Injured Burn Patient				5a. CONTRACT NUMBER W81XWH-12-1-0550	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Peter Morris, James Holmes, Brad Freeman, Bruce Cairns, Michael Berry, L. Doug Case E-Mail: pemorris@wakehealth.edu				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Wake Forest University Health Sciences, Winston-Salem, NC Washington University, St. Louis, MO University of North Carolina, Chapel Hill, NC				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT This is a multicenter, randomized controlled trial to determine whether early ICU rehabilitation, for Burn Intensive Care Unit (BICU) patients will decrease hospital length of stay. 50 subjects will be randomized at each of three sites for a total of 150 subjects. The study has completed all regulatory requirements, completed site protocol developments and has begun to enroll patients. The goal enrollment minimums are an average of 2.5 patients enrolled per month, per site; 7.5 patients enrolled per month, across the study. This study will increase understanding of the effect of rehabilitation on ICU Burn patients, through ultrasound and strength assessments of muscles, performed at study entry (ultrasound), ICU & Hospital discharge and at 3, 6 and 12 months (ultrasound & strength assessments) post-enrollment. Functional testing with Short Physical Performance Battery (SPPB) and Health Related Quality of Life (HRQoL) testing will determine if standardized early rehab improves functional performance, quality of life and employment status. Accomplishments Year #1: Database build, design web entry case report forms, site training; finalized IRB consent forms and began enrollment, 3 subjects to date, with outpatient follow-up 3, 6, and 12 month sessions planned. Accomplishments Year #2: Active Study with enrollment of study subjects at all three sites. Completion of training of ultrasound techniques at all three sites. Out-patient follow-up has begun for those enrolled within the first year of the study.					
15. SUBJECT TERMS Critical Injury, Burn Patient, Rehabilitation Therapy					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	39	19b. TELEPHONE NUMBER (include area code)

Table of Contents

	<u>Page</u>
Introduction.....	4
Keywords.....	4
Overall Project Summary.....	4
Key Research Accomplishments.....	5
Conclusion.....	5
Publications, Abstracts, and Presentations	6
Inventions, Patents, and Licenses.....	7
Reportable Outcomes.....	7
Other Achievements.....	7
References.....	7
Appendices.....	7

Introduction:

This project is funded in order to conduct a multicenter, randomized controlled trial to determine whether early ICU rehabilitation, for Burn Intensive Care Unit (BICU) patients, will decrease hospital length of stay. Fifty subjects will be randomized at each of three sites for a total of 150 subjects. Study start-up was initiated in Year #1 and all sites are now enrolling patients. Sixteen study subjects have been enrolled. Out-patient visits during the post-enrollment, post-hospital discharge phase of the study have begun and outpatient phase of testing has been conducted. As well, phone follow-ups at the end of second and third years post-enrollment, remain planned. The sites are now committed to increasing enrollment goals above planned – and will average of 2.5 patients enrolled per month, per site; 7.5 patients enrolled per month or greater, across the study, to compensate for the time spent on regulatory work prior to initiation of subject enrollment.

This study will proceed and allow for greater understanding of the effect of rehabilitation on the pathophysiology of ICU Burn patients. Through ultrasound and strength assessments of muscles, performed at study entry (ultrasound), ICU & Hospital discharge and at 3, 6 and 12 months post-enrollment, the study will determine if Standardized Rehabilitation Therapy (SRT) decreases loss of bicep and quadricep muscle mass, architecture change and prevents strength loss. The protocol contains functional testing with the Short Physical Performance Battery (SPPB) and Health Related Quality of Life (HRQoL) testing with SF-36 and Burn Specific Health Scale (BSHS-B). These tools will be administered to determine if SRT improves functional performance, quality of life and employment status.

Keywords:

Burn Injury, Critical Care, Intensive Care, Standardized Rehabilitation Therapy

Overall Project Summary:

Monthly study conferences have been active with discussion across PI's and coordinators. These conferences have led to recommendations to address the gap between targeted and actual enrollment volume within the study. These conversations were detailed and examined possibilities at each site to identify improvements in process. Additionally and more importantly, it was noted that the burn population at these sites were in receipt of pre-hospital management that had changed since the original proposal. The PI's then elected to put into action a plan that would increase enrollment by recognizing this factor of pre-hospital burn patient management and have now addressed this practice more appropriately within the inclusion criteria. Changing the inclusion criteria to eliminate mechanical ventilation as the key criterion, has allowed focus on the severity of the burn & its sequelae. The change has been reviewed within the PI group, local IRBs and the HRPO and will not materially alter the study other than to increase enrollment. Given that pre-hospital procedural changes throughout the US have occurred with respect to intubation over the past few years, resulting in fewer intubated patients arriving in EDs, intubation status is now not a significant indicator of burn severity. Burn size, which determines ICU admission, is the key determinant for burn severity. Thus, being able to enroll any patient warranting Burn ICU admission will allow us to focus on the severely burned patient & demonstrate the benefits of early aggressive ICU rehab, which has yet to be demonstrated in the peer-reviewed literature.

Patient screening and patient enrollment:

Patient screening commenced and has been ongoing at all three sites. Through August 2014, 345 individuals have been screened for this study since May, 2013.

Patient review:

To date, sixteen study subjects have been enrolled. Appropriate timelines were adhered to in regards to study enrollment windows. Appropriate execution of study inclusion and exclusion rules was conducted. Each study subject had an appropriate study IRB consent form, with appropriate dated signatures, obtained prior to randomization. Randomization procedures were engaged and functioned without difficulty.

To date both study arms have been engaged with study subjects. Standardized rehabilitation therapy and usual care were delivered to study subjects. Success was achieved in the delivery of multiple intervention arm rehabilitation sessions including delivery of resistance training exercise with Therabands. Blinded exercise physiologists have conducted the strength and functional assessments according to protocol.

Our publications have targeted the study's standardization of ultrasound images of muscles in an ICU population.

Key Research Accomplishments:

- All IRB and HRPO obligations have been met
- All subcontract sites have working relationships with Wake Forest to receive study payments
- Electronic secure remote entry database is functioning
- Ultrasound muscle imaging protocol has been completed allowing for standardized images across sites
- Study screening is ongoing
- Study subject enrollment and randomization is ongoing
- Intervention sessions are being delivered according to protocol
- Ultrasound images are being obtained
- Strength and Functional Assessments are being obtained
- Plans for augmentation of study subject enrollment has been discussed by investigators with recognition of need to surpass initial proposed enrollment timelines for timely study completion

Conclusion:

To date, the investigators conclude that the work performed will allow the study hypotheses to be tested sufficiently. Current hypotheses are that standardizing rehabilitation for Intensive Care Unit Burn patients will:

- 1) shorten hospital length of stay in burn patients.
- 2) prevent loss in muscle size and loss of architecture during critical illness of severe burns.

- 3) improve objective strength, functional measures, and quality of life at 3, 6, 12, 24, and 36 months post-enrollment.

The investigators conclude that the systems built and implemented for this study will allow for the conduct of a multicenter, randomized controlled trial which will determine whether early ICU, standardized rehabilitation therapy for BICU with blinded strength assessments decreases hospital length of stay. Fifty (50) subjects will be randomized at each of three sites, for a total of 150 subjects.

Although study startup was prolonged, subject enrollment is now active. Enrollment criteria have been adjusted to better optimize the volume of screened patients who become enrolled. Sixteen study patients have been enrolled. The study subjects will receive three in-person visits during the 1st year post-enrollment and will be followed by phone follow-up at end of 2nd and 3rd years post-enrollment. Enrollment strategies will be engaged to enhance the initial grant's proposal of 2.5 patients enrolled per month, per site; 7.5 patients enrolled per month, across the study.

Ultrasound and strength assessments of muscles have been successfully performed at study entry, ICU & Hospital discharge. These will continue at 3, 6 and 12 months post-enrollment. These observations will be critical for the study to determine if standardized rehabilitation therapy decreases loss of biceps and quadriceps size, architecture change and blunts strength loss.

Lastly, performance of functional testing with the Short Physical Performance Battery (SPPB) and HRQoL testing with SF- 36 and Burn Specific Health Scale (BSHS-B) will determine if SRT improves functional performance, quality of life and employment status.

Publications, Abstracts, and Presentations:

Our group along with colleagues in Melbourne, Australia has examined the reliability of ultrasound image assessments within the ICU population. This work has now been compiled into a manuscript and has been submitted for publication (see appendix).

a. Manuscripts:

1. Lay Press:
2. Peer-Reviewed Scientific Journals:

A Sarwal, S Parry, M Berry, F Hsu, M Lewis, N Justus, P Morris, L Denehy, S Berney, S Dhar, M Cartwright. Inter-observer Reliability of Quantitative Muscle Ultrasound Analysis in the Critically Ill Population. JUSM. 2014;

3. Invited Articles:
4. Abstracts:

- i. S Dhar, A Sarwal, D Files, J Hauser, P Campbell, C Thompson, P Morris. Muscle Ultrasound & Short Physical Performance Battery (SPPB) to Predict ICU Functional Outcomes. Crit Care Med. 2013; 41 (Suppl 1):A89
- ii. A Sarwal, K Bringolf, M Cartwright, F Hsu, M Berry, D Bowton, S Dhar, P Morris. Early Serial Changes on Diaphragm Muscle Ultrasound May Predict Duration of Mechanical Ventilation. Crit Care Med. 2013; 41 (Suppl 1):A89-90

- iii. Parry SM, El-Ansary D, Berney S, Koopman R, Cartwright MS, Morris PE, Hilton A, Sarwal A, Denehy L. The Role Of Neuromuscular Ultrasound Imaging And Relationship To Muscle Strength And Physical Function In The Critically Ill Population. ATS, San Diego, May 2014.
- iv. S Parry, A Sarwal, M Cartwright, P Morris, L Denehy. Reliability of Quantitative Muscle Ultrasound Analysis in the Critically Ill Population. Crit Care Med. 2013; 41 (Suppl 1):A108

b. Presentations: by Peter Morris, MD (PI)

- 1. November 2013, Bethesda MD, National Annual Meeting of the United States Critical Illness and Injury Trials Group, New Paradigms for Early ICU Rehab
- 2. January 2014, Society of Critical Care Medicine, United States Critical Illness and Injury Trials Group meeting, Update on Early ICU Rehabilitation
- 3. March 2014, Duke Children's Optimization of Mechanical Ventilation Conference, Update On ICU Rehab Techniques, Raleigh, NC
- 4. May 2014, American Thoracic Society, The Future of Early ICU Rehabilitation. San Diego, CA

Inventions, Patents, and Licenses:

Nothing to report

Reportable Outcomes:

Nothing to report

Other Achievements:

Nothing to report

Appendices:

Manuscript and abstract publications

MUSCLE ULTRASOUND & SHORT PHYSICAL PERFORMANCE BATTERY (SPPB) TO PREDICT ICU FUNCTIONAL OUTCOMES

Sanjay Dhar, Aarti Sarwal, DC Files, M Berry, J Hauser, P Campbell, C Thompson, Peter Morris;

Introduction: Difficulties predicting ICU patient functional outcomes include inability to adequately assess the patient's true functional status just prior to acute illness or injury. ICU admission muscle ultrasound may serve to provide characterization of pre-ICU muscle architecture quality. Ultrasound image alterations such as an increase in mean gray scale value may indicate increased muscle echogenicity; similar changes have been reported in muscular dystrophies and chronic inflammatory myopathies. Such a non-invasive, safe and easily obtained image, may provide investigators and clinicians with a novel tool to better judge an ICU patient's degree of pre-morbid muscle function. Such insight may allow for better understanding of the contribution of a poor pre-morbid functional status on post-ICU functional status. This independent baseline contribution may be measureable in the future and be additive to muscle dysfunction attributable to the duration of ICU exposure. **Methods:** 16 patients enrolled into the NIH supported "Standardized Rehabilitation Therapy for ICU Patients with Acute Respiratory Failure" study (5R01NR011186-05), underwent standard study protocol hospital discharge SPPB administration. In addition, these 16 underwent biceps muscle ultrasound at study entry. B-mode image acquisition was obtained using an M-turbo system (Sonosite) with a 6–15 MHz transducer. Biceps image obtained 10 cm prox to antecubital fossa. Image J was used to obtain gray scale data, per our previous publication, Cartwright, MuscleNerv '13; 47:255. Baseline demographics and hosp days were collected. Pearson correlation coefficients were performed using biceps gray scale mean and SPPB score. **Results:** Mean baseline values of the 16 patients included age: 51 years (12.6 std); APACHE III: 70, (18 std). All were mechanically ventilated at study entry. Baseline Biceps gray scale mean 54.4 (21.1 std). The 16 patients demonstrated a mean Hosp duration of 11.4 days, (6.6 std). Hospital discharge SPPB mean score was 4.6 (3.8 std). [SPPB scoring tool Range 0–12; 12 = highest functioning score]. Pearson correlation between baseline biceps mean and hospital discharge SPPB was a (-)0.54, $p=0.029$. **Conclusions:** ICU admission images with increased echogenicity on muscle ultrasound imaging as determined by Image J analysis demonstrated an association with a low SPPB functional assessment at hospital discharge. These pilot data suggest that there may be an association between the demonstration of an abnormal baseline biceps ultrasound image and the patient's eventual hospital discharge functional status as determined by the SPPB score. This pilot study holds potential to lend novel pre-ICU muscle architecture information as it pertains to assessing pre-morbid muscle architecture's role in determining pre-ICU, baseline functionality. Future study of such image assessments may allow investigators to become better informed of a patient's pre-morbid functional status; lack of pre-morbid functional status is a current barrier in determining what role acute illness plays in ICU & Hospital discharge functional status. With further study, such baseline assessments may prove useful to target ICU patients who may especially benefit from a patient-specific prescription of an intense ICU rehab strategy, due to the ICU admission identification of pre-ICU low functional status characteristics on biceps muscle ultrasound imaging.

EARLY SERIAL CHANGES ON DIAPHRAGM MUSCLE ULTRASOUND MAY PREDICT DURATION OF MECHANICAL VENTILATION.

Sarwal, Aarti; Bringolf, Karamie; Cartwright, Michael; Hsu, Fang-chi; Berry, M; Bowton, David; Dhar, Sanjay; Morris, Peter

Introduction: Diaphragm muscle weakness from disuse atrophy, critical care weakness or ventilator induced muscle dysfunction may contribute to weaning failure in patients on prolonged mechanical ventilation (MV). Difficulty in weaning patients accounts for a large proportion of time spent in the intensive care unit (ICU). Patients who fail extubation and get re-intubated have higher mortality even after controlling for the presence of co-morbid conditions and generalized severity of illness. Many studies have shown evidence of atrophy and degeneration in diaphragm muscle using electromyography and muscle biopsy in subjects on MV. These tests are not feasible or practical for serial assessments of changes in diaphragm muscle in critically ill patients. Simple, clinical tools are needed that can identify these muscle changes early on in critically ill patients to better predict patients at risk of extubation failure so we can develop and tailor early, patient-specific strategies for successful weaning. We performed this pilot study to assess serial changes in diaphragm muscles of patients on MV and determine whether these muscle changes predict extubation failure and length of MV.

Methods: We recruited 14 mechanically ventilated patients admitted to the NeuroICU and Trauma ICU at Wake Forest Baptist Medical Center, Winston Salem, NC over a 3 month period. The study protocol was approved by the Institutional Review Board and informed consent was obtained on all patients. Baseline ultrasound exam was done to image diaphragm in right mid axillary line within 72 hours of ICU admission. Serial exams were repeated daily until patients were extubated or transferred out of the ICU, for a maximum of 7 ultrasound studies. Patients were excluded if they had a prior history of MV, chest tube or thoracoabdominal surgery. Images were analyzed using Image J for diaphragm muscle thickness at end expiration and echodensity using gray scale histogram. Linear regression analysis was done to determine if serial changes in each muscle value predicted days on MV after adjusting for age, Body Mass Index & cumulative fluid balance.

Results: 14 ventilated patients (10 males) had mean age 55.2 +/- 20 years, BMI 27.79 +/- 6.22. Baseline mean diaphragm thickness within 72 hours of ICU admission was 2.70 +/- 0.61 mm & did not exclusively predict the days of MV. 11 patients with serial scans had 5.73 +/- 4 days on MV, 3.45 days between 1st and last scan and fluid balance 5.31 L +/- 8.8L on average. We observed increase in diaphragm thickness and decrease in echodensity over time. On regression analysis, serial change in diaphragm echodensity was a significant predictor of total number of days on MV after adjusting for age, BMI, and fluid balance on day of last USG ($p = 0.0285$).

Conclusions: Our preliminary data shows the feasibility of performing quantitative serial measurements on diaphragm muscle ultrasound in critically ill patients and suggests that early serial changes in the muscle parameters may predict duration of MV. The early serial changes detected by ultrasound suggest a potential benefit of proactive early therapies designed to preserve respiratory muscle architecture to reduce days on MV and prevent extubation failure. Small sample size precludes definite interpretation of trends seen in our study. Larger studies are needed to define parameters of serial diaphragm muscle changes on ultrasound in patients on MV. Muscle ultrasound has potential benefit as a non-invasive tool to assess serial changes in respiratory muscles in ICU patients on MV. It can be used to test optimal ventilator modes and weaning strategies resulting in reduced ICU length of stay and optimize long term outcomes in patient on MV.

The Role Of Neuromuscular Ultrasound Imaging In The Critically Ill Population And Relationship To Muscle Strength And Physical Function

S. M. Parry^{1, 2}, D. El-Ansary¹, S. C. Berney³, R. Koopman⁴, M. S. Cartwright⁵, P. E. Morris⁶, A. Hilton⁷, A. Sarwal⁵, L. Denehy¹,

¹Department of Physiotherapy, The University of Melbourne, Parkville, VIC, Australia, ^{2, 3}Austin Health, Melbourne, VIC, Australia, ⁴Department of Physiology, The University of Melbourne, Parkville, VIC, Australia, ⁵Department of Neurology, Wake Forest School of Medicine, Winston-Salem, NC, ⁶Wake Forest University Health Sciences, Winston Salem, NC, ⁷Intensive Care Unit, Austin Health, Melbourne, VIC, Australia

Corresponding author's email: selina.parry@austin.org.au

RATIONALE: Early and accurate diagnosis of intensive care acquired weakness is challenging in the clinical setting. Currently muscle strength testing is the predominant tool used for diagnosing this weakness. Ultrasonography is a promising non-volitional method for tracking musculoskeletal changes from admission. No studies have previously examined the relationship between ultrasound measurements of muscle structure and muscle strength or physical function in the critically ill population. The aims were 1) to assess the relationship between muscle structure, muscle strength, and physical function in patients who are critically ill; 2) to assess difference in muscle thickness and gray scale changes between baseline and ICU discharge.

METHODS: Adults ventilated for >48 hours were included and underwent serial muscle ultrasound measurements for thickness and grayscale analysis of the quadriceps muscle complex (rectus femoris (RF), vastus intermedius (VI) and vastus lateralis (VL) from admission to ICU discharge. Measurements of muscle strength and physical function were performed on awakening and at ICU discharge using Medical Research Council sum-score, physical function in intensive care test-scored and highest functional level attained. Statistical analysis included paired t-tests and Pearson correlations.

RESULTS: Twenty-two patients were included. Baseline demographics are shown in Table 1. There were no significant relationships between baseline RF and VI thickness and muscle strength on awakening or at ICU discharge. RF thickness at ICU discharge had a moderate positive correlation with discharge measures of physical function ($r=0.574$, $p=0.020$) and highest functional level ($r=0.711$, $p=0.002$). There was a moderate positive correlation between VL thickness and ICU discharge muscle strength ($r=0.508$, $p=0.05$) and physical function ($r=0.509$, $p=0.044$). Baseline RF grayscale analysis had a moderate negative correlation with discharge muscle strength ($r=-0.611$, $p=0.015$). Discharge measures of VI grayscale analysis had a moderate negative correlation with discharge measures of physical function ($r=-0.602$, $p=0.014$) and also diagnosis of ICU-acquired weakness ($r=-0.599$, $p=0.018$). There were significant differences in muscle thickness for RF (mean difference=20%, $p=0.002$) and VI muscles (mean difference=11%, $p=0.018$) between baseline and ICU discharge. There was also a significant difference over time in RF grayscale standard deviation scores (mean difference=-6.90, $p=0.020$).

CONCLUSIONS: Ultrasonography is a promising bedside tool for detecting changes in muscle architecture in the critically ill population. Ultrasonographic measures of muscle structure, muscle strength and physical function are moderately related in the critically ill population.

Table 1: Baseline demographics

Variable	Mean (SD) or n, %
Age, years	56.2 (17.8)
Gender, male	13 (59.1)
Diagnosis Category	
- Medical	14 (63.6)
- Surgical	8 (36.4)
Time from admission to baseline ultrasound measures, hours	51.0 [20.9 - 69.9] ¹
APACHE II Score	23.4 (8.1)
Incidence of ICU-AW	9 (40.9)
MV time, hours	172.1 [112.3 - 284.1] ¹
Time to awakening, days	10.0 (4.3)
ICU LOS, days	11.5 [9.0 - 21.5] ¹
Hospital LOS, days	22.0 [12.8 - 41.3] ¹
Discharge Destination - Home	11 (50.0)

Abbreviations: ICU-AW, intensive care unit acquired weakness; ICU, intensive care unit; LOS, length of stay; MV, mechanical ventilation.

¹ Median [interquartile range]

This abstract is funded by: None

Am J Respir Crit Care Med 189;2014:A4505

Internet address: www.atsjournals.org

Online Abstracts Issue

Inter-observer Reliability of Quantitative Muscle Ultrasound

Analysis in the Critically Ill Population

(Original Research)

Running title: Inter-observer reliability in ICU

Aarti Sarwal, MD;^{1,2} Selina M. Parry, PT;^{3,4} Michael J. Berry, PhD;⁵ Fang-Chi Hsu, PhD;⁶

Marc T. Lewis;⁵ Nicholas W. Justus;⁵ Peter E. Morris, MD;⁷ Linda Denehy, PhD;⁴

Sue Berney, PhD;³ Sanjay Dhar, MD;⁷ Michael S. Cartwright, MD, MS¹

From the Departments of Neurology¹ and Anesthesiology (Section on Critical Care Medicine),² Wake Forest School of Medicine, Winston-Salem, North Carolina, USA; Physiotherapy,³ Austin Health, Melbourne, Australia; University of Melbourne,⁴ Melbourne, Australia; Health and Exercise Science,⁵ Wake Forest University, Winston-Salem, North Carolina, USA; Biostatistical Sciences,⁶ Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA; Internal Medicine (Section on Critical Care Medicine),⁷ Wake Forest School of Medicine, Winston-Salem, North Carolina, USA

Corresponding Author: Aarti Sarwal, Department of Neurology, Wake Forest School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157 Telephone: 336-716-4101
E-mail: asarwal@wakehealth.edu

ABSTRACT

Objective: There is growing interest in the use of quantitative high-resolution neuromuscular ultrasound imaging to evaluate skeletal muscles in patients with critical illness. There is currently considerable methodological variability in the measurement technique of quantitative muscle analysis. Reliability of muscle parameters using different measurement techniques and assessor expertise level has not been examined in patients with critical illness. The primary objective of this study was to determine the inter-observer reliability of quantitative ultrasound measurement analyses (thickness and echogenicity) between assessors of different expertise levels and using different techniques of selecting region of interest.

Materials and Methods: Cross-sectional observational study in a Neurocritical care and a mixed surgical-medical intensive care unit from two tertiary referral hospitals.

Results: Twenty diaphragm and 20 quadriceps images were evaluated. Images were obtained using standardized imaging acquisition techniques. Quantitative ultrasound measurements included muscle thickness and echogenicity analysis (either by trace or square techniques). All images were analyzed twice independently by four assessors of differing expertise levels. Excellent inter-observer reliability was obtained for all measurement techniques regardless of expertise level (Intra-class Correlation Coefficient (ICC) >0.75 for all comparisons). There was less variability between assessors for echogenicity values when square technique was used for quadriceps muscle and the trace technique for the diaphragm.

Conclusions: Excellent inter-observer reliability exists regardless of expertise level for quantitative analysis of muscle parameters on ultrasound in the critically ill population. Based on

43 these findings, it is recommended that echogenicity analysis be performed using square
44 technique for quadriceps and trace technique for diaphragm.

45

46 Key words: critical care weakness; grayscale analysis; muscle, ultrasound

INTRODUCTION

Traditional techniques of muscle evaluation including manual muscle testing, electromyography, nerve conduction studies, and muscle biopsy have proven challenging in the routine, clinical, and research studies of intensive care unit-acquired weakness (ICU-AW) because of the need for trained staff, patient cooperation, or invasive nature of the study.¹⁻³ Muscle biopsy studies have shown that ultrastructural changes occur in skeletal muscle echotexture in ICU-AW.⁴ There is growing interest in the development of a noninvasive imaging tool for muscle function and structure with high clinical utility to evaluate these changes in the intensive care unit (ICU) setting.

The evolution of high resolution ultrasound imaging of the muscle has allowed us to use quantitative muscle parameters to distinguish muscle pathologies such as atrophy, dystrophy, and sarcopenia.^{5,6} Quantitative muscle parameters on B-mode ultrasound imaging – muscle thickness and gray scale echogenicity – have been proposed to reflect changes in muscle echotexture.^{7,8} Muscle thickness on ultrasound images has been validated against other imaging modalities such as computed tomography and magnetic resonance imaging as well as direct measurements on dissected cadavers.⁹⁻¹¹ Mean echogenicity value of a muscle is calculated by performing grayscale analysis of image pixels using the histogram feature of image processing software such as Image J or Photoshop© (Figure 1). All the pixels in the selected area of the muscle are categorized on a gray scale configuration. This has been shown to correlate with muscle pathology on biopsy.^{8,12} Normative values of both thickness and echogenicity have been reported in healthy volunteers and neuromuscular disease states.¹³⁻¹⁵

Recently, muscle ultrasound imaging has attracted significant attention in assessing skeletal muscle changes seen in critical illness, sepsis, and mechanical ventilation.^{4,16} Ultrasound is non-invasive, safe and readily available in the ICU with no risk of radiation; hence, it is an ideal modality to study ICU-AW. Reliability of ultrasound image acquisition by operators of various experience levels has been shown to be high but we do not know if experience plays a role in facilitating more reliable image analysis if a standardized protocol of image analysis is used.¹⁷⁻¹⁹ There is also considerable methodological variability in quantitative muscle analysis in this emerging area. Most of these reliability studies have been performed in images obtained from healthy volunteers or patients with neuromuscular diseases in outpatient settings.²⁰ There is an urgent need to evaluate the reliability and differences in quantitative ultrasonography measurements using different techniques to determine if differences exist in results obtained by different assessor expertise levels in the ICU setting.

This is the first study to examine the inter-observer reliability of quantitative ultrasound image analysis in patients in the ICU setting. The primary objective of this study was to determine the inter-observer reliability of quantitative ultrasound measures of thickness and echogenicity in both the diaphragm and quadriceps muscles for novice and experienced operators. The secondary objective was to determine if there was a difference between quantitative echogenicity techniques used. Initial results of this analysis were presented at the Society of Critical Care Medicine Congress, San Francisco, CA in January 2014.²¹ The Guidelines for Reporting Reliability and Agreement Studies were followed for preparation of the manuscript.^{22,23}

MATERIALS AND METHODS

Study design and setting

This was a cross-sectional observational study examining reliability of muscle ultrasound imaging analysis in previously acquired muscle images from critically ill patients admitted to the ICU in two different hospital settings. Diaphragm images were acquired in patients admitted to the Neurocritical care unit at Wake Forest School of Medicine, Winston-Salem, NC, USA. Quadriceps images were acquired in patients admitted to a mixed surgical-medical ICU at Austin Health, Melbourne, Australia. Images were acquired as a part of ongoing clinical studies at both centers with relevant approval by the local institutional review board (IRB). Written informed consent was obtained for images acquired at Wake Forest per IRB requirements. Human Research Ethics Committee at Austin Health waived the need for informed consent.

Image acquisition

All ultrasound images were obtained with a standardized protocol of transducer placement, anatomic landmarks, and patient position for acquisition. The same experienced assessor acquired all images included for analysis (diaphragm – AS; quadriceps – SMP) with depth, gain, compression and sonographic settings kept constant between subjects. Rectus femoris images were obtained using an 8.5-MHz linear-array transducer (Voluson E BT09 Ultrasound, GE Healthcare, Yokogawa Medical Systems Ltd., Japan) at 5 cm depth on the right lower limb. The transducer was placed perpendicular to the long axis of the thigh on its anterior surface, two-thirds of the distance from the anterior superior iliac spine to the superior patellar border. The specific orientation of the probe was chosen as this was the only orientation that

could be clinically ~~the most~~ be standardized without needing additional equipment in the critical care setting. Also, our previous work in critically ill patients suggested that the cross-sectional ultrasonographic view might be more sensitive to changes in muscle echotexture in ICU population.¹ Imaging was completed in supine position with the leg supported in passive extension and neutral rotation.^{1,24} Diaphragm images were obtained on the right side via an intercostal approach using a 6-15 MHz Transducer (M-Turbo Ultrasound System, SonoSite, Inc., USA) placed in the sagittal plane with patients in supine position and the head of bed elevated at 30 degrees. Images were captured in the right mid-axillary line to obtain a sagittal image of diaphragm muscle at the zone of apposition at end of expiration with the transducer always perpendicular to the diaphragm muscle.²⁵

Study procedures

Assessors

Four assessors were involved in this study. Two assessors were experienced ultrasonographers (AS-neurocritical care physician and SMP-Senior ICU Physiotherapist) with ~~more than~~ one year of experience in muscle ultrasound acquisition and quantitative muscle imaging analysis. Currently, there is no formal specified credentialing or certification process for muscle ultrasonography. Our assessors were considered experienced based on their background formal training in neuromuscular physiology, muscle ultrasound techniques and established clinical and research experience of acquiring and analyzing muscle ultrasound images using Image J.^{1,24-37} The other two assessors, considered novice users (ML and NJ), were exercise physiology students with knowledge of muscle anatomy but no prior ~~ultrasound~~ exposure to ultrasonography

or image analysis software. Both received a total of three hours of training over the course of two sessions on ultrasound and quantitative muscle analysis using Image J software, a public domain, Java-based image processing program developed at the National Institute of Health.³⁸ The training was done by one of the experienced assessors (AS) and included demonstration of ultrasound image acquisition with hands on training using a volunteer to familiarize them with muscle anatomy on ultrasonography. Then they were given instructions on identifying anatomy and artifacts on ultrasound images and on using Image J. A written handout was given that outlined muscle anatomy, major artifacts, Image J troubleshooting and detailed instructions on measuring each parameter with pictorial examples. They were given a series of sample images to analyze at home. Their competency was tested by having them demonstrate each measurement to the instructor for each method multiple times at the end of session 1, beginning and end of session 2.

For the actual study, the images were de-identified by a neutral person (MJB) and then randomly assigned to each assessor for image analysis. The assessors were blinded in between the analysis by time delay of 1 month.

Image analysis

Twenty diaphragm and 20 quadriceps images obtained in critically ill patients during different stages of ICU stay were randomly selected for quantitative analysis. Operator selecting images was blinded to clinical characteristics of patients from which these images were obtained. These images were reviewed in a random order by each rater following a standardized protocol described below. Repeat analysis was done by all assessors blinded to previous results. All four

assessors performed measurements independently and had no access to the results from the other assessors or their own previous assessments thereby ensuring blinding of results from one another. All assessors repeated the echogenicity, and thickness measurements twice and the mean value was used in statistical analyses. All images were analyzed with Image J using a standardized protocol for measurement of thickness and echogenicity for both diaphragm and quadriceps musculature.³⁸

Thickness

Muscle thickness was measured as maximum distance between the pleural and peritoneal margins of the diaphragm muscle.³⁵ Quadriceps thickness was measured as the maximum thickness of rectus femoris/vastus intermedius complex from femur to subcutaneous tissue.¹

Echogenicity

Muscle echogenicity was measured using the gray scale analysis histogram feature on Image J.³⁹ Mean echogenicity was measured in an area of the muscle outlined using the “trace” method as well as the “square” technique (Figure 1). Using the trace method, the assessor highlighted all visible muscle area excluding artifacts and analyzed the enclosed area for echogenicity analysis.³⁹ For the square technique, a standard square area (20×20 pixels for diaphragm and 100×100 pixels for rectus femoris) was chosen within the muscle margins that best represented the muscle bulk.¹ Pixel length was used for defining the area due to different pixel density at various depths used to produce ultrasound images. For the diaphragm, echogenicity was measured on the muscular portion beyond aponeurosis whenever possible.

When analyzing echogenicity of quadriceps, only rectus femoris was used for echogenicity analysis when anatomically visible otherwise the representative area of the rectus femoris/vastus intermedius complex was chosen. For the square method, if the defined pixel measures (20×20 for diaphragm and 100×100 for rectus femoris) were found to be larger than cross section of the muscle, the largest square that stayed within the anatomic boundaries of the muscle was chosen.

Statistical analysis

All statistical analyses were performed using SPSS for MacIntosh Statistical software package (Mac SPSS Version 21, IBM, New York, NY, USA). One Sample Kolmogorov-Smirnov test was used to test for normality of the data. Mean (\pm SD) of mean measures for muscle thickness and echogenicity for the quadriceps and diaphragm muscles for the four raters were calculated. One way ANOVA was performed to compare differences among the mean values obtained for each muscle parameter by different raters (Tables 1 and 2).

Intra-Class Correlation Coefficient (ICC) values were calculated using the mean of measures and the ICC model (2.1) two-way random effects for inter-observer reliability (Table 3). An ICC is measured on a scale of 0 to 1; 1 represents perfect reliability with no measurement error, whereas 0 indicates no association.²³ An ICC >0.75 is considered to have good to excellent reliability.^{40,41} Paired t-tests were performed to determine mean differences and 95% confidence between mean echogenicity using square versus trace methods. The level of significance was set at 0.05. Bland-Altman (BA) plots were examined to provide a visual representation of the agreement between assessors for thickness and echogenicity measurements for both the

diaphragm and quadriceps muscles and to examine if there is a systematic error present (Figure 2).

Sample Size Calculation

To look for a minimum ICC value of 0.75 of inter-observer reliability, we estimated a priori that a minimum of 12 participants were required for each muscle analyzed to achieve 80% power with significance level set at 0.05. To ensure that enough images were available for statistical analysis, we included 20 diaphragm and 20 quadriceps images in this study.^{40,42}

RESULTS

Mean (\pm SD) of mean measures for muscle thickness and echogenicity for the diaphragm and quadriceps muscles for the four raters are shown in Tables 1 and 2. One sample Kolmogorov-Smirnov test was used to test for normality of the data and no significant differences were present between the observed distributions and normal distributions. No significant differences were found among the four raters in diaphragm muscle thickness or echogenicity as determined by the trace or square method or in quadriceps muscle thickness or echogenicity as determined by the trace or square method.

Inter-observer reliability results between the two novice raters and inter-observer reliability results between the two experienced raters are shown (Table 3). Overall, there was very good inter-observer reliability between all assessors for measurement of muscle thickness and echogenicity regardless of technique (trace or square) for both diaphragm and quadriceps muscles with ICC values ranging from 0.84 to 0.99. Experienced raters showed similar ICCs for muscle thickness and echogenicity for the quadriceps and diaphragm muscles. Diaphragm thickness showed the least agreement and consistency of all muscle parameters between the experienced raters, although the correlation values were still high. For the novice raters, overall ICCs for thickness and echogenicity via the square method were lower in the diaphragm compared to the quadriceps. Reliability of quadriceps thickness and echogenicity obtained from the square method was excellent (>0.90) for both the novice and experienced raters. Between the two novice assessors, the trace method for echogenicity analysis for quadriceps muscles showed wider limits of agreements, which were statistically significant ($p<0.01$) with a mean difference of -6.05 mm.

When comparing the two different methods of echogenicity analyses, there was a significant difference in mean echogenicity values between the square and trace method for measurement of echogenicity in each muscle. For the quadriceps, the mean difference between square and trace was 9.26 (95% CI: 5.73-12.79; $p < 0.001$ and for the diaphragm, the mean difference between square and trace was 1.14 (95% CI: 0.23-2.05; $p = 0.02$).

Bland-Altman plots for agreement between all assessors for thickness and echogenicity (square and trace) for diaphragm and quadriceps muscles are shown in Figure 2. Bland-Altman analysis demonstrated wider limits of agreements in novice raters for all quantitative measurements – thickness and echogenicity, for both diaphragm and quadriceps muscles when compared to experienced assessors (Figure 2).

DISCUSSION

Our study shows that quantitative analysis of muscle parameters on ultrasound images acquired using a clinically reproducible standardized image acquisition and image analysis protocol is reliable and feasible in critically ill patients. The findings support further use and research related to standardized quantitative ultrasound analysis in the critical care setting.

Many previous studies have shown inter-observer and intra-observer reliability for acquisition of muscle ultrasound images in healthy volunteers.^{23,43-46} A recent study evaluated the inter-observer reliability of gray scale density in large peripheral limb muscles of 19 boys with Duchene muscular dystrophy and 21 healthy control boys and found high reliability with similar ICCs.¹⁸ We undertook this study to evaluate reliability of quantitative neuromuscular ultrasound imaging in the critically ill population that may have different and unique muscle characteristics. Diaphragm and quadriceps muscle were chosen because they have significant differences in anatomy and acquisition characteristics. The diaphragm muscle is relatively thin (mean reported thickness 3.3 mm) and challenging to identify especially on post-acquisition review of static B-mode imaging.^{25,47} In comparison, the quadriceps is a relatively thick muscle (mean reported thickness 44 mm) and thus potentially easier to image and identify.¹³ Ultrasound image analysis depends on knowledge of anatomy as well as a working understanding of ultrasound artifacts. This study was designed to assess if analysts with variable training background and knowledge of anatomy and ultrasound were able to measure quantitative ultrasound parameters with a certain degree of consistency and agreement since prevalent research and clinical studies have variable training and backgrounds of personnel performing ultrasonography. Observer dependency has been criticized as major factor impeding widespread

use of ultrasonography in clinical and research settings. This study assessed the reliability of observations made by novice users since many clinical research designs involve training of personnel with limited exposure or experience in muscle ultrasonography.

Our study design has a fundamental difference from published protocols on measuring echogenicity that focus on reporting accuracy of echogenicity. We formulated the design of the study to allow easy reproducibility of image acquisition and analysis protocols in further research and clinical studies. Evidence suggests that probe orientation needs to be perpendicular to the muscle fibers to have the greatest echogenicity and a slight tilt or pressure can change the echogenicity significantly.^{10,48-50} The published protocols use or suggest use of a device or a robot arm that ensures uniform force and angle targets during the time of image acquisition once fiber directions is ascertained with visualization.⁵¹ While we did consider anisotropy and possibility of using such a device, this device is not a routine part of clinical ultrasonography and its requirement during imaging may limit the widespread clinical use of quantitative ultrasonography in assessing muscle pathology. It is also very difficult in the ICU to standardize the probe directions to be perpendicular to the fibers as patients may have confounding factors like atrophy, edema, invasive devices, or pathology that may impair fiber visualization. Our goal was to keep the image acquisition protocol reproducible and practically useful in clinical settings so that further studies can investigate the utility of muscle echogenicity in assessing muscle weakness in the critical care settings and investigate its correlation to clinical outcomes. Many authors have highlighted the need for validation and reliability studies in critically ill patients and our study is first in series of such an attempt.⁵² Use of stringent image acquisition criteria that

may be challenging to reproduce in routine clinical settings would hamper and restrict further investigations and widespread clinical use of muscle echogenicity parameter.

Currently echogenicity measurements are not available as part of standard ultrasonography software. Computer assisted measurements have to be done post acquisition using image analysis software to measure gray scale density like Adobe Photoshop© and Image J.³⁸ We chose ultrasonography machines most commonly used in routine clinical settings in the critical care units .The choice of an image analysis software was based on the fact it was designed by NIH for scientific work ,is available for free download along with easy instructions on use and allows batch processing hence can analyze a batch of images quickly . These features allow feasibility of clinical use until inbuilt ultrasonographic software becomes available for such measurements. By showing reliability of measurements under clinically reproducible settings with novice users, our hope is to allow larger investigational studies that focus on finding clinically useful markers of muscle pathology that can be used to guide diagnosis or interventions. Muscle ultrasonography has a wide range of applicability in the critical care settings due to its portable non invasive nature.⁵³⁻⁶¹ Muscle echogenicity is being investigated for use to assess change in a patient's nutritional status that be a target of nutritional therapies.^{7,52,62} If a clinical protocol using muscle echogenicity can identify patients at higher risk of muscle changes upon admission or onset of weakness during ICU stay, interventions like early and more aggressive physical therapy , nutritional therapies or functional muscle stimulation have a potential of reducing the extent of weakness and ultimately reducing length of stay and improving functional outcomes.^{24,36,37,63} In addition, diaphragm muscle ultrasonography has

potential of identifying patients at risk of extubation failure and such patients can be targeted for specific weaning protocols that strengthen respiratory muscle.^{21,64-67}

Overall, excellent correlations between novice and experienced analysts for all ultrasound measures in our study reflects that quantitative muscle analysis can be learned with focused training on muscle anatomy and ultrasound techniques despite having no prior training in ultrasonography or computer assisted image analysis negating the widely held belief that learning ultrasonography has a steep learning curve that is recently being challenged.¹⁸

Thickness values had high reproducibility overall for both muscles. Diaphragm thickness showed the least agreement and consistency due to the challenge of identifying it post processing. The muscle is very small and may be hard to distinguish individually from the intercostal muscle.

Thickness measurement is also challenged by layers of high echogenicity produced by intertwining layers of diaphragm tendons that may confuse the observer about the actual span of the diaphragm thickness.³⁵ Reliability of quantitative echogenicity analysis has not been previously described for the diaphragm muscle. In our study, the best inter-observer reliability of mean echogenicity values was obtained using the trace technique for the diaphragm. Use of the square method for diaphragm analyses may have made it more difficult to identify the muscular portion of the diaphragm and hence not as reliable. However, the square technique had more consistent results for the quadriceps muscle amongst both novice and experienced users in our cohort. Part of this discrepancy might be a result of the variability generated as the assessor was arbitrarily able to choose to include the rectus femoris/vastus intermedius complex when anatomy of the rectus femoris was not clearly defined. This may include the inter-muscular

connective tissue, fascia and blood vessels. Restricting to a predefined square for analysis focussed most area on muscular bulk that reflected in higher consistency.

Mean values of thickness and echogenicity on our patients varied slightly from normative values reported in healthy volunteers emphasizing the need to assess normative ranges for quantitative ultrasound parameters in critically ill patients who might have varied pre-existing level of fitness medical comorbidity or chronic illness all of which affect baseline muscle parameters.¹³ Mean thickness values for diaphragm at end expiration in our cohort ranged from 2.2-2.7 mm. Mean diaphragm thickness has been reported in literature to be 3.3 mm with 1.5 mm as lower limit of normal resting thickness.⁴⁷ These values are reported in outpatient setting evaluations in normal subjects done at standardized lung volumes. Most critically ill patients are unable to cooperate to allow measurements at a set lung volume. Mean quadriceps thickness reported in normal, healthy volunteers ranges between 36-47 mm while our population had mean thickness between 17.1-8.2 mm.¹³ Our images were obtained in critically ill patients who might have had some degree of muscle weakness and atrophy attributed to comorbidities like congestive heart failure, chronic obstructive pulmonary disease, or critical care weakness. This could explain the thinner quadriceps in our cohort.

Our study has several limitations. We have limited data on critically ill patients on clinically significant differences between normal and abnormal muscle thickness and echogenicity values. Our study also lacks correlations of quantitative muscle parameters to clinical markers of muscle strength and structure. The clinical significance of the limits of agreement values seen for muscle parameters in our study also needs to be explored. One study analyzing seven intubated patients showed decrease in diaphragm thickness by 6% per day of

mechanical ventilation.⁶⁸ Whether such small changes can be reliably detected given variability of diaphragm measurements needs to be investigated.

Conclusions

Quantitative muscle analysis on ultrasound images has excellent inter-observer reliability with agreement and consistency amongst experienced analysts and novice users after structured formal training. While we reiterate the need for a standardized protocol for image acquisition and analyses, our results suggest that measures of echotexture using gray scale analysis might not be interchangeable between square and trace techniques. We suggest that studies using quantitative ultrasound should report on the exact methodology (square or trace) utilized in muscle analysis as well as expertise and number of different assessors used. Future research needs to be focussed on correlating clinical measures of muscle function to quantitative muscle parameters such as thickness and echogenicity values. High resolution quantitative muscle ultrasound holds great potential in clinical and research evaluation of muscle changes in the ICU setting.

Acknowledgments:

The authors acknowledge ultrasound support from the Center of Medical Ultrasound at Wake Forest School of Medicine, Winston Salem, North Carolina, USA, and the Departments of Physiotherapy and Intensive Care at Austin Health, Melbourne, Australia. The team is also grateful for the faculty and staff working in the intensive care units at each institute for providing excellent care to our patients and supporting clinical research and to Adela Larimore for assistant with manuscript preparation and submission.

Financial Support:

Aarti Sarwal received administrative support to conduct the study as a research scholar with the Translational Science Institute funded by NCRR/NIH M01 RR007122.

Sanjay Dhar received honorarium from American College of Chest Physicians and American Thoracic Society for teaching ultrasound courses.

Selina M Parry has funding through the National Health and Medical Research Council Dora Lush Scholarship (#103923), Society of Critical Care Medicine, Austin Medical Research Foundation and the Intensive Care Foundation Australia.

Linda Denehy has grants to study intensive care rehabilitation and neuromuscular ultrasound from the Austin Medical Research Foundation and the Intensive Care Foundation Australia.

Sue Berney has funding through the National Health and Medical Research Council Early Career Fellowship (#1037283), Austin Medical Research Foundation and the Intensive Care Foundation Australia

383 Peter E. Morris and Michael J Berry have funding through NIH/NINR/NHLBI R01NR011186-

384 01

385 Michael S Cartwright has funding through NIH/NINDS 1K23NS062892.

386

387 **Competing Interests:** None of the authors have any competing interests to declare

388

References

48. Kwah LK, Pinto RZ, Diong J, Herbert RD. Reliability and validity of ultrasound measurements of muscle fascicle length and pennation in humans: a systematic review. *J Appl Physiol (1985)*. 2013;114(6):761-9.
49. Herbert RD, Gandevia SC. Changes in pennation with joint angle and muscle torque: in vivo measurements in human brachialis muscle. *J Physiol*. 1995;484(Pt 2):523-32.
50. Ishida H, Watanabe S. Influence of inward pressure of the transducer on lateral abdominal muscle thickness during ultrasound imaging. *J Orthop Sports Phys Ther*. 2012;42(9):815-8.
51. Harris-Love MO, Monfaredi R, Ismail C, Blackman MR, Cleary K. Quantitative ultrasound: measurement considerations for the assessment of muscular dystrophy and sarcopenia. *Front Aging Neurosci*. 2014;6:172.
52. Mourtzakis M, Wischmeyer P. Bedside ultrasound measurement of skeletal muscle. *Curr Opin Clin Nutr Metab Care*. 2014;17(5):389-95.
53. Berry MJ, Morris PE. Early exercise rehabilitation of muscle weakness in acute respiratory failure patients. *Exerc Sport Sci Rev*. 2013;41(4):208-15.
54. Carlucci A, Ceriana P, Prinianakis G, Fanfulla F, Colombo R, Nava S. Determinants of weaning success in patients with prolonged mechanical ventilation. *Crit Care*. 2009;13(3):R97.
55. Caruso P. Ventilator-induced diaphragmatic dysfunction: keep working. *Crit Care Med*. 2005;33(12):2852-3.

- 410 56. Gottesman E, McCool FD. Ultrasound evaluation of the paralyzed diaphragm. *Am J*
411 *Respir Crit Care Med.* 1997;155(5):1570-4.
- 412 57. Hermans G, Agten A, Testelmans D, Decramer M, Gayan-Ramirez G. Increased duration
413 of mechanical ventilation is associated with decreased diaphragmatic force: a prospective
414 observational study. *Crit Care.* 2010;14(4):R127.
- 415 58. Matamis D, Soilemezi E, Tsagourias M, et al. Sonographic evaluation of the diaphragm
416 in critically ill patients. Technique and clinical applications. *Intensive Care Med.*
417 2013;39(5):801-10.
- 418 59. Morris PE, Goad A, Thompson C, et al. Early intensive care unit mobility therapy in the
419 treatment of acute respiratory failure. *Crit Care Med.* 2008;36(8):2238-43.
- 420 60. Puthuchear Z, Harridge S, Hart N. Skeletal muscle dysfunction in critical care: wasting,
421 weakness, and rehabilitation strategies. *Crit Care Med.* 2010;38(10 Suppl):S676-82.
- 422 61. Summerhill EM, El-Sameed YA, Glidden TJ, McCool FD. Monitoring recovery from
423 diaphragm paralysis with ultrasound. *Chest.* 2008;133:737-43.
- 424 62. Puthuchear ZA, Rawal J, McPhail M, et al. Acute skeletal muscle wasting in critical
425 illness. *JAMA.* 2013;310(15):1591-600.
- 426 63. Moisey LL, Mourtzakis M, Cotton BA, et al. Skeletal muscle predicts ventilator-free
427 days, ICU-free days, and mortality in elderly ICU patients. *Crit Care.* 2013;17(5):R206.
- 428 64. Kim SH, Na S, Choi JS, Na SH, Shin S, Koh SO. An evaluation of diaphragmatic
429 movement by M-mode sonography as a predictor of pulmonary dysfunction after upper
430 abdominal surgery. *Anesth Analg.* 2010;110(5):1349-54.

- 431 65. Kim WY, Suh HJ, Hong SB, Koh Y, Lim CM. Diaphragm dysfunction assessed by
432 ultrasonography: influence on weaning from mechanical ventilation. *Crit Care Med*.
433 2011;39(12):2627-30.
- 434 66. Petrof BJ, Jaber S, Matecki S. Ventilator-induced diaphragmatic dysfunction. *Curr Opin*
435 *Crit Care*. 2010;16(1):19-25.
- 436 67. Doorduyn J, van Hees HW, van der Hoeven JG, Heunks LM. Monitoring of the
437 respiratory muscles in the critically ill. *Am J Respir Crit Care Med*. 2013;187(1):20-7.
- 438 68. Grosu HB, Lee YI, Lee J, Eden E, Eikermann M, Rose KM. Diaphragm muscle thinning
439 in patients who are mechanically ventilated. *Chest* 2012; 142:1455-1460.

Table 1. Quantitative Parameters for Quadriceps Muscle. Each box represents mean of scores (Standard Deviation)

	Rater A	Rater B	Rater C	Rater D	p-value
Muscle Thickness	1.82 (0.75)	1.80 (0.70)	1.80 (0.80)	1.71 (0.68)	0.96
Square Echogenicity	77.0 (21.2)	74.5 (20.8)	79.30 (24.8)	79.9 (24.5)	0.87
Trace Echogenicity	65.0 (17.1)	71.0 (20.1)	68.4 (17.2)	69.4 (19.4)	0.77

A between measures ANOVA was used to test for significant differences among the 4 raters. No significant differences were found among (between) the 4 raters for quadriceps muscle thickness, square echogenicity or trace echogenicity.

Table 2: Quantitative Parameters for Diaphragm Muscle*

	Rater A	Rater B	Rater C	Rater D	p-value
Muscle					
Thickness	0.22 (0.08)	0.23 (0.08)	0.26 (0.08)	0.25 (0.08)	0.54
Square					
Echogenicity	43.6 (13.4)	45.2 (14.4)	48.8 (15.4)	48.7 (13.8)	0.57
Trace					
Echogenicity	43.4 (13.5)	44.9 (14.7)	46.9 (13.6)	46.7 (14.3)	0.86

* Each box represents mean of scores (Standard Deviation).

A between measures ANOVA was used to test for significant differences among the 4 raters. No significant differences were found among (between) the 4 raters for diaphragm muscle thickness, square echogenicity or trace echogenicity.

Table 3: Inter-Observer Reliability Assessment

	Novice Inter-observer Reliability		Experienced Inter-observer Reliability	
	ICC (95% CI)		ICC (95% CI)	
	<u>Agreement</u>	<u>Consistency</u>	<u>Agreement</u>	<u>Consistency</u>
Quadriceps				
Muscle thickness	0.968 (0.921-0.987)	0.966 (0.917-0.987)	0.913 (0.796-0.965)	0.917 (0.802-0.966)
Square				
echogenicity	0.902 (0.774-0.960)	0.904 (0.775-0.961)	0.993 (0.984-0.997)	0.993 (0.983-0.997)
Trace				
echogenicity	0.844 (0.537-0.943)	0.883 (0.729-0.952)	0.973 (0.935-0.989)	0.973 (0.934-0.989)
Diaphragm				
Muscle thickness	0.851 (0.665-0.938)	0.859 (0.678-0.942)	0.868 (0.683-0.940)	0.882 (0.727-0.952)
Square				
echogenicity	0.851 (0.667-0.938)	0.851 (0.661-0.938)	0.964 (0.911-0.986)	0.962 (0.907-0.955)
Trace				
echogenicity	0.899 (0.767-0.959)	0.899 (0.764-0.959)	0.984 (0.959-0.993)	0.983 (0.958-0.993)
ICC, intra-class correlation coefficient; 95% CI, 95 percent confidence intervals				

Figure Legends

Figure 1: Echogenicity analysis of the diaphragm muscle using (upper left) a predefined square area of 20×20 pixels (square technique) and (upper right) tracing the anatomical boundary of the muscle (trace technique). Echogenicity analysis of rectus femoris muscle using (lower left) a predefined square area of 100×100 pixels (square technique) and (lower right) tracing the anatomical boundary of the muscle (trace technique).

Figure 2: Bland-Altman analysis plots for agreement between assessors for diaphragm and quadriceps muscle. The x-axis shows the mean of two values as assessed by each assessor. The y-axis shows the difference between means of these values for each assessor. The three horizontal lines parallel to x-axis represent the length of agreement (LOA+), the mean difference and LOA- for that muscle parameter.